COMPLETED

768-J-01

L 2.71/5-2: 92-14

National Longitudinal Surveys

U.S. Department of Labor
Bureau of Labor Statistics

Discussion Paper

95-025874 [illegible]

The National Longitudinal Surveys (NLS) program supports many studies designed to increase understanding of the labor market and methods of data collection. The *Discussion Papers* series incorporates some of the research coming out of the program. Many of the papers in this series have been delivered as final reports commissioned as part of an extramural grant program, under which grants are awarded through a competitive process. The series is intended to circulate the findings to interested readers within and outside the Bureau of Labor Statistics.

Persons interested in obtaining a copy of any *Discussion Paper*, other NLS reports, or general information about the NLS program and surveys should contact the Bureau of Labor Statistics, Office of Economic Research, Washington, DC 20212-0001 (202-606-7405).

Material in this publication is in the public domain and, with appropriate credit, may be reproduced without permission.

Opinions and conclusions expressed in this document are those of the author(s) and do not necessarily represent an official position of the Bureau of Labor Statistics or the U.S. Department of Labor.

NLS Discussion Papers

Self-Selection and Internal Migration in the United States

George J. Borjas
Stephen G. Bronars
Stephen J. Trejo

March 1990

Report: NLS 92-14

Self-Selection and Internal Migration in the United States

George J. Borjas
University of California at San Diego, and Unicon

Stephen G. Bronars
University of Texas, and Unicon

Stephen J. Trejo
University of California at Santa Barbara, and Unicon

March 1990

This final paper was funded by the U.S. Department of Labor, Bureau of Labor Statistics under contract number J-9-J-7-0093. Opinions stated in this paper do not necessarily represent the official position or policy of the U.S. Department of Labor.

SELF-SELECTION AND INTERNAL MIGRATION IN THE UNITED STATES

George J. Borjas, Stephen G. Bronars, and Stephen J. Trejo

March 1990

This project was funded by the U.S. Department of Labor Bureau of Laoor Statistics under Contract Number J-9-J-7-0093. Opinions stated in this document do not necessarily represent the official position or policy of The U.S. Departrment of Labor.

SELF-SELECTION AND INTERNAL MIGRATION IN THE UNITED STATES

George J. Borjas, Stephen G. Bronars, and Stephen J. Trejo

Executive summary

The population of the United States is highly mobile. Since the 1960s, approximately three percent of the population moves across state lines in any given year, and 10 percent of the population moves across state lines in a five-year period. In view of the falling fertility rates of American women, this extensive mobility implies that migration has become an increasingly important source of demographic change in the various regions and a major determinant of concurrent changes in regional economic growth.

Internal migration rates are especially high among the young. In the National Longitudinal Survey of Youth, over half of the sample are currently (i.e., in 1986) residing in a state other than their state of birth, and about 40 percent are living in a different state than they were at age 14 (see Table 1). The data also indicate that approximately 8 percent of white young men and 6 percent of black and Hispanic young men move across state boundaries in any given year. Our study presents a theoretical and empirical analysis of the internal migration of young workers in the United States.

Guided by income-maximizing models, previous research has focused on "explaining" the size and direction of migration flows across areas. Generally, these studies conclude that persons migrate from low-income regions to high-income regions, and that high migration costs deter interregional mobility. We extend this framework by analyzing not only the size and direction of migration flows, but also their skill composition. In addition to regional income differentials, our model stresses the importance

TABLE 1

INTERSTATE MIGRATION RATES IN NATIONAL LONGITUDINAL SURVEYS OF YOUTH

	Whites		Blacks		Hispanics	
	Males	Females	Males	Females	Males	Females
% Living in Different State Than at Birth	52.4	50.1	40.8	43.7	57.6	57.9
% Living in Different State Than at Age 14	41.3	38.8	29.9	33.4	37.1	36.8
% Living in Different State Than in 1979	32.1	28.5	23.3	25.4	20.6	21.1
% Moving Across State Lines Per Year:						
1979-1980	8.6	8.4	6.8	7.8	6.6	5.7
1980-1981	8.2	7.8	5.8	6.6	6.8	5.7
1981-1982	9.8	7.7	5.3	8.4	8.4	7.6
1982-1983	8.7	8.1	6.6	7.3	4.9	7.2
1983-1984	5.9	4.6	3.9	4.3	3.1	2.7
1984-1985	5.6	5.2	4.6	5.0	2.9	3.1
1985-1986	8.8	7.4	5.6	5.8	4.5	5.9

√

of regional differences in the returns to skills. Because the impact of internal migration on economic conditions depends not only on how many persons move, but also on which persons move, we believe that these questions are as important as those that have dominated the literature.

In this framework, each region is characterized by an earnings distribution, and persons must decide which among the various earnings opportunities is best for them (net of migration costs). The potential income which could be earned by a particular worker in a particular region depends on the mean level of earnings offered in the region, as well as on the rate of return paid to the specific skills possessed by the worker. Some regions offer higher mean incomes than other regions, and, similarly, some regions offer higher prices for skills than other regions.

Income-maximizing behavior generates a well-defined equilibrium sorting in which differences in rates of return to skills among regions are the main determinants of the skill composition of internal migration flows. In particular, regions that pay higher returns to skills attract more skilled workers than regions that pay lower returns. This result implies that highly skilled workers currently residing in a region that offers a relatively low payoff for those skills, or that unskilled workers currently residing in a region that penalizes their lack of skills, are the most mismatched spatially, and therefore these workers have the greatest incentive to move. In an important sense, our theoretical approach extends earlier models by emphasizing the role of prices in allocating people among regions.

To test our theory, we use the 1979-1986 waves of the National

Longitudinal Survey of Youth. Respondents are between the ages of 14 and 22 at the time of the first interview, and the subsequent annual surveys provide a detailed history of each individual's labor force activity and geographic mobility. In our empirical work, we assume that the region of residence at age 14 is exogenously determined by parental location decisions. Because optimal location decisions for parents and children need not coincide, a spatial mismatch may arise for youths. Our empirical analysis, therefore, focuses on the study of migration patterns after age 14, as youths relocate themselves in order to correct these initial mismatches. Further, we focus on migration across state boundaries--rather than across regional or county boundaries--throughout the study.

To analyze the skill composition of the migrant flow, we use alternative measures of a worker's skills, including educational attainment, wages (both with and without adjustment for differences in demographic characteristics), and the worker's score on the Armed Forces Qualification Test. The theory predicts that highly skilled workers are most likely to leave states with low rates of return to skill, because highly skilled workers are most mismatched in these states. We estimate mover/non-mover probit models where mover status is determined by race and sex dummy variables and by skills.

The results indicate that, in general, migration rates are higher for the more skilled. However, as predicted by the theory, the data also indicate that an increase in skills has a larger impact on the migration propensity in states which offer relatively small payoffs to skill. For example, a one standard deviation increase in years of education raises the probability that a worker leaves a state with a low return to skills by about 5.3 percent, while the same change in skills has a trivial impact (.8% percent) on the migration propensities of individuals in high return states. Therefore, it appears that high ability workers are migrating out of states with low returns to skills and low ability workers are migrating out of states with high returns to skills.

Our theoretical model also predicts that skilled workers will move to states with high returns to skills. We test this implication by viewing the change in the rate of return to skills in the state of residence between age 14 and 1986 as a choice variable. We find that this variable is positively related to skill endowments. Put differently, skills are an important determinant of the direction of internal migration flows. Moreover, these effects are economically important. A one standard deviation increase in years of education increases the rate of return to skills in the destination state by about 25 percent.

The empirical analysis, therefore, provides important insights into the internal migration process. Individuals are more likely to move the greater the mismatch between skill endowments and the returns to skills. Moreover, the direction of internal migration flows are guided by comparative advantage. Highly skilled workers tend to migrate to states that pay a larger reward to labor market skills, while unskilled workers are likely to end up in states where earnings are relatively insensitive to skills. Overall, the empirical evidence suggests that internal migration plays an important role in determining the equilibrium sorting of skills across states.

I. SELF-SELECTION AND INTERNAL MIGRATION IN THE UNITED STATES

SELF–SELECTION AND INTERNAL MIGRATION IN THE UNITED STATES

George J. Borjas, Stephen G. Bronars, and Stephen J. Trejo*

I. Introduction

The population of the United States is highly mobile. Since the 1960s, approximately three percent of the population moves across state lines in any given year, and 10 percent of the population moves across state lines in a five–year period (U.S. Department of Commerce, annual). In view of the falling fertility rates of American women, this extensive mobility has become an increasingly important source of demographic change in the various regions and a major determinant of concurrent changes in regional economic growth.

The theoretical and empirical study of internal migration has a long history in labor economics (Greenwood, 1975, 1985). The substantive importance of this topic arises because internal migration helps equilibrate economic conditions across labor markets in a competitive economy. The study of the size, direction, composition, and economic impact of mobility flows across labor markets can be used to test the extent to which the neoclassical model provides a reasonable description of labor market conditions in the United States (Topel, 1986).

Guided by the income–maximizing models of Hicks (1932) and Sjaastad (1962), early empirical research focused on "explaining" the size and direction of migration flows across areas, and determining why certain groups of individuals, such as the highly educated, are more likely to migrate than others (Bowles, 1970; Greenwood, 1969; Schwartz, 1973; Vanderkamp, 1971). During the mid–1970s, economists began to focus on the role of the family in the migration decision (DaVanzo, 1976; Mincer, 1978; Polachek and Horvath, 1977; and Sandell, 1977). These studies yield a number of interesting empirical results and theoretical insights, and suggest that Becker's (1981) approach for modeling economic behavior within the family provides a useful framework for the analysis. Despite the central role played by internal migration in the operation of a competitive economy, we believe it is fair to conclude that the study of internal migration has not been at the forefront of research in mainstream labor economics in the past decade.[1]

The recent disinterest in the study of internal migration flows is not symptomatic of a more general indifference in the study of geographic mobility. After all, the analysis of international migration flourished in recent years (Borjas, 1987, 1990; Chiswick, 1978; Freeman and Abowd, 1990). This burgeoning literature suggests that combining the Hicks–Sjaastad income–maximizing approach with the insights provided by Roy's (1951) self–selection model generates new substantive insights about the immigration process, and helps resolve many of the empirical puzzles in international migration data.

This paper presents an application of this more general approach to the analysis of internal migration flows in the United States. We argue that the Hicks–Sjaastad migration model provides a much too restrictive framework for the study of internal migration. Its key predictions are that persons migrate from low–income regions to high–income regions and that higher mobility costs deter migration. In turn, these predictions help focus the empirical work almost exclusively on "explaining" the size and direction of population flows across regions. Although the data generally support these predictions, there are many other interesting questions left unaddressed by the model and by the empirical studies.

The Hicks–Sjaastad model is restrictive because (for given migration costs) it focuses solely on the fact that mean income levels differ across regions, and these income differentials generate one–way migration flows. The theoretical approach suggested by the Roy model stresses regional differences in the returns to skills (in addition to the income differentials). We will show that these price differentials are the main determinants of the skill composition of the population flow across regions. Regions that pay higher returns to skills attract more skilled workers than regions that pay lower returns. Our analysis thus raises fundamental questions about the skill composition of internal migrants. Because the impact of internal migration on economic conditions depends not only on how many persons move, but also on which persons move, we believe that these questions are as important as those that have dominated the literature.

Of course, the literature that estimates the returns to migration has been well aware of the fact that migrants are self-selected. The development of econometric techniques to account for selection bias (Heckman, 1979) encouraged a number of applications of this methodology to the analysis of migrant earnings. Studies by Nakosteen and Zimmer (1980), Hunt and Kau (1985), and Robinson and Tomes (1982) report the standard selectivity-corrected earnings functions in the mover and stayer samples. The evidence in these papers is inconclusive because of the sensitivity of the methodology, and also because the results are not couched within the equilibrium sorting predicted by the economic theory of selection.

This paper presents a theoretical and empirical analysis of the role self-selection plays in determining the size, direction, and skill composition of internal migration flows. Within the conceptual framework provided by the Roy model, we present an empirical analysis of internal migration flows using the National Longitudinal Survey of Youth. As shown in Table 1, internal migration is very prevalent among these young men and women. Over 50 percent of the persons in the sample are currently (i.e., in 1986) residing in a different state than at the time of birth, and about 40 percent are living in a different state than at age 14. The data also indicate that approximately 8 percent of white young men and 6 percent of black and Hispanic young men move across state boundaries in any given year.

Our empirical analysis indicates that interstate differences in the rewards to skills are a major determinant of internal migration flows. Generally, persons whose skills are most mismatched with the reward structure offered by their current state of residence are the persons most likely to migrate out of that state. Moreover, the data indicate that these persons tend to relocate in states which offer higher rewards for their particular skills. The empirical evidence, therefore, implies that the Roy model provides a useful framework for analyzing internal migration flows in the United States.

II. Theory

Consider a country partitioned into k distinct geographic regions, indexed by i=1,...,k. To simplify the exposition, we initially assume that there are no costs of relocating across regions. After birth, therefore, individuals compare their earnings opportunities in the various regions and move to the region that maximizes their earnings. This approach provides a simple and intuitive description of the equilibrium skill sorting generated by endogenous migration flows. Further, we assume that the initial distributions of individual skills are the same for all regions. At the time of birth, therefore, individuals are randomly allocated across regions in terms of their skills. The population log earnings distribution in region i is given by:

$$\log w_i = \mu_i + v_i, \quad i=1,\ldots,k \tag{1}$$

where μ_i is the mean income that would be observed in region i in the absence of any internal migration, and v_i is a random variable with mean zero and variance σ_i^2 that measures person–specific deviations from mean income in region i.

The assumption that the initial skill distributions are identical across regions makes the income distributions in (1) independent of initial conditions: the same earnings determination process in region i applies to all persons regardless of their region of birth. This assumption ignores the possibility that persons born in region i are distinctly different, on average, from persons born in region j, and hence the income distributions should also be subscripted for region of birth. The main results of our model are unaffected by this complication. Hence we maintain the simplifying assumption that the initial allocation of skills does not affect the shape of the population income distributions in any of the regions.[2] However, because of regional differences in the level of resources, physical capital, and aggregate economic conditions, mean incomes μ_i and the distribution of the random variables v_i will vary across regions. Therefore, there will be regional differences in the shapes of the income distributions, even in the absence of differences in the initial skill distributions of workers.

An individual chooses to reside in region j whenever:

$$\log w_j > \max_{r \neq j} [\log w_r] \qquad (2)$$

As is well known (Heckman and Honore, 1987; Roy, 1951), the characteristics of the sorting generated by the decision rule in (2) cannot be described without additional restrictions on the distribution of the random variables $v_1,...,v_k$. A simplifying assumption which allows a complete characterization of the equilibrium sorting is that individual earnings are perfectly correlated across regions, so that $\mathrm{Corr}(v_i, v_j) = 1$ for all i,j. The population income distribution of region i can then be written as:

$$\log w_i = \mu_i + \eta_i v, \quad i=1,...,k \qquad (1')$$

This formulation of the regional differences in earnings opportunities implies that the same random variable v determines individual earnings in the various regions (up to a factor of proportionality). It is instructive to think of v as indexing the worker's ability or skills. In effect, equation (1′) assumes that the earnings determination process can be characterized by a one–factor model of ability. The factor of proportionality η_i can be thought of as the factor–loading parameter, or more generally as the "rate of return" to skills in region i. It is useful to arbitrarily order the regions such that they they are ranked in terms of η with $\eta_1 < \ldots < \eta_k$. We also assume that v is a continuous random variable with a range defined over the real number line. This is the only restriction we make on the density of skills in the population.

Although the assumption that earnings are perfectly correlated across regions is quite strong (for it implies that the ranking of individuals by skill level is the same in all regions), it enables us to derive a number of testable implications from a multi–region selection model. Furthermore, this framework may provide a reasonably accurate representation of earnings opportunities across regions in the United States, given the relatively strong regional similarities in culture, law, and economic institutions.

It is worth stressing that we have not addressed the issue of whether the random variable v is observed or unobserved to the researcher. From an individual's point of view, this distinction is irrelevant. Individuals sort themselves across regions on the basis of all their skills, not just those that happen to be econometrically convenient. By using a one–factor model of ability, we assume that the relative prices of all skills are the same across regions, so that the composite commodity theorem allows us to focus on a single good that is being "sold" across regions. It is not difficult to analyze the migration decision within the context of a multi–factor model of ability, but we do not pursue this generalization because it detracts from the essential points that we make in this paper.

Using equations (1') and (2), it is easy to show that income–maximizing behavior implies that the equilibrium sorting of skills across the k regions is given by:

$$\text{Choose region 1:} \quad -\infty < v < \min_{i=2,\ldots,k} \left[\frac{\mu_1-\mu_i}{\eta_i-\eta_1}\right] \tag{3a}$$

$$\underset{(1 < j < k)}{\text{Choose region } j\text{:}} \quad \max_{i=1,\ldots,j-1} \left[\frac{\mu_i-\mu_j}{\eta_j-\eta_i}\right] < v < \min_{i=j+1,\ldots,k} \left[\frac{\mu_j-\mu_i}{\eta_i-\eta_j}\right] \tag{3b}$$

$$\text{Choose region k:} \quad \max_{i=1,\ldots,k-1} \left[\frac{\mu_i-\mu_k}{\eta_k-\eta_i}\right] < v < \infty \tag{3c}$$

In view of the assumption that v takes on values from minus to plus infinity, equations (3a) and (3c) imply that some persons will necessarily reside in the "extreme" regions 1 and k (i.e., the regions with the minimum and maximum η). It is less apparent, however, that some persons choose to locate themselves in any other region. In particular, it may well be the case that for some parameters values, equation (3b) will not be satisfied by region j, and hence no workers will choose to reside there. After all, it would seem that persons with positive v's have the most to gain by migrating to the region with the highest η, while persons with negative v's have the most to gain by moving to the region with the lowest η.

Because unpopulated regions are of no interest empirically, we restrict our attention to "interior solutions", i.e., to regions where some individuals choose to reside. Equation (3b) implies that a necessary condition for region j to be inhabited is $\frac{\mu_{j-1}-\mu_j}{\eta_j-\eta_{j-1}} < \frac{\mu_j-\mu_{j+1}}{\eta_{j+1}-\eta_j}$. This restriction can be rewritten as:

$$\mu_j > \frac{(\eta_{j+1}-\eta_j)\mu_{j-1} + (\eta_j-\eta_{j-1})\mu_{j+1}}{(\eta_{j+1}-\eta_{j-1})} \tag{4}$$

Equation (4) defines the Existence Condition that mean incomes in region j must satisfy in order to attract and retain a population. We assume that the Existence Condition is satisfied by all j (j=2,...,k−1). This assumption greatly simplifies the characterization of the equilibrium sorting. In particular, the repeated use of the Existence Condition to make pairwise comparisons of the arguments in the min(·) and max(·) terms in (3) implies that the sorting of skills across regions can be written as:

Back to
prep for
renumbering